

PROJEKTMESSE „ELEKTRONISCHE LEXIKOGRAFIE“

von Carolin Müller-Spitzer und Marie Vachková

Begleitend zur Jahrestagung wurde dieses Jahr eine Projektmesse zur elektronischen Lexikografie veranstaltet. Hierzu wurden Projekte eingeladen, die zum einen im weitesten Sinne im Bereich der elektronischen Lexikografie beheimatet sind, die aber in ihrer Projektarbeit insbesondere auch Themenkomplexe berühren, die in den Kontext des Themas der diesjährigen Jahrestagung – „Sprachliches Wissen zwischen Lexikon und Grammatik“ – passten.

1. Als erster Teilnehmer ist das „Algemeen Nederlands Wordenboek“, das Allgemeine Niederländische Wörterbuch, zu nennen. Das ANW ist die Neuerarbeitung eines allgemeinen, einsprachigen Wörterbuchs gezielt für das elektronische Medium. Das Besondere hier ist u. a. der Ansatz einer formalisierten Bedeutungsbeschreibung durch sog. Semagramme, durch die interessante und neuartige erweiterte Suchen im elektronischen Wörterbuch ermöglicht werden sollen.

2. Das Prager Projekt, „Das Große akademische Wörterbuch Deutsch-Tschechisch“, stellte sich vor allem mit der neu entwickelten Online-Datenbank vor. Dieses Projekt reflektiert in seiner begleitenden Forschung die neuen Anforderungen der Korpuslinguistik und setzt korpusanalytische Methoden um, die am IDS im Programmbereich Korpuslinguistik erarbeitet werden. Das Thema der lexikalischen Beziehungen (*near synonyms*) steht dabei im Vordergrund.



Projektmesse im Rosengarten

3. Im Projekt „Deutscher Wortschatz der Universität Leipzig“ finden sich u. a. automatisch ermittelte Kollokationen sowie Ober- und Unterbegriffe; über die zugrunde liegenden Methoden und Arbeitsweisen konnte man sich an diesem Stand informieren.
4. Das „Digitale Wörterbuch der deutschen Sprache“ basiert auf dem „Wörterbuch der deutschen Gegenwartssprache“, kurz WDG, mittlerweile sind aber viele Funktionen und Angaben zu dieser Grundsubstanz hinzugekommen. In der neuen Version des DWDS, die im Internet in einer Beta-Version bereits angesehen werden kann, sind zudem verschiedene Sichten auf die unterschiedlichen Substanzen (Wörterbücher, DWDS-Kerncorpus etc.) vorgesehen.
5. Ganz anderen Problemen und Herausforderungen hat sich das „Korpusbasierte, elektronische Wörterbuch Deutsche Gebärdensprache – Deutsch“ zu stellen. In diesem Projekt spielen unterschiedliche Fragen eine Rolle, z. B. wie ein gebärdensprachliches Korpus aufgebaut wird, welche Datenschutzprobleme es im Vergleich zur Schriftsprache gibt oder wie man im Kontext von Gebärdensprache mit Varietäten umgeht.
6. „GermaNet – ein lexikalisch semantisches Wortnetz“ ist ein computerlexikografisch ausgerichtetes Projekt, welches mehr auf die automatische Sprachverarbeitung als auf menschliche Benutzer ausgerichtet ist. GermaNet reiht sich in eine ganze Serie auf das internationale WordNet folgende nationale WordNet-Vorhaben ein, ist mittlerweile integriert in das EuroWordNet und bot auf der Projektmesse eine interessante Ergänzung zu anderen eher korpuslinguistisch ausgerichteten lexikografischen Projekten.
7. Im Dortmunder „HyTex-Projekt“ wurden „Werkzeuge zur Exploration von GermaNet“ entwickelt, z. B. eine grafische Sicht auf GermaNet. Da für GermaNet im akademischen Kontext mittlerweile kostenfrei Lizenzen vergeben werden, werden auch diese Werkzeuge an Bedeutung gewinnen.
8. Das multilinguale elektronische Wörterbuch zur Fußballsprache, das „Kicktionary“, beruht auf einem framebasierten semantischen Ansatz. Hier kann man sich das sprachliche Inventar unterschiedlicher „Scenes“ wie z. B. „Shot“, den Schuss, in vielen Sprachen ansehen. Das Kicktionary war daher nicht nur in Hinblick auf die naheende Fußball-WM einen Besuch wert, sondern auch aufgrund seines semantischen Ansatzes.
9. Wiederum in den Kontext der Frametheorie gliedert sich das „Saarbrücken Lexical Semantics Acquisition Project“ (SALSA) ein. Ziel des Projektes ist es, ein großes framebasiertes Lexikon für das Deutsche bereitzustellen, mit möglichst reichhaltigen semantischen und syntaktischen Annotationen. Ähnlich wie GermaNet, ist auch SALSA als Ressource zunächst für linguistische und computerlinguistische Forschungen sowie für NLP (Natural Language Processing)-Anwendungen gemacht.
10. Im Kontext des eldit-Projektes, eines digitalen bilingualen Online-Wörterbuchs Deutsch-Italienisch, ist die Dissertation zu „Kognitiv Salienten Semantischen Relationen“ angesiedelt. Dieses Dissertationsprojekt soll im Ergebnis eine Hilfe bieten bei der Erweiterung von eldit im Bereich der Anzeige von semantischen Relationen zu einem gesuchten Lexikoneintrag (in der Lerner-Sprache).



Projektmesse im Rosengarten

11. Als nächstes folgen zwei Projekte, die sich mit Valenz beschäftigen und von daher ganz eng in den Kontext der Jahrestagung passten. Zunächst die „Erlangen Valency Patternbank“, in der auf Basis des „Valency Dictionary of English“ eine Datenbank aufgebaut wird, in der Konstruktionen und Muster nach unterschiedlichen Kriterien durchsucht und sortiert werden können.
12. Das am IDS beheimatete Projekt „E-VALBU“ ist eine für das elektronische Medium aufbereitete Version des gedruckten Valenzwörterbuchs deutscher Verben. Das v. a. in der DaF-Welt sehr gut angenommene Wörterbuch wurde für die Online-Version gezielt überarbeitet und eröffnet durch zahlreiche erweiterte Suchen interessante Abfragemöglichkeiten.
13. Auch eng in den Kontext der Jahrestagung durch die Beschäftigung mit Konstruktionen passten die zwei Projekte der Abteilung Lexik des IDS, die Wortverbindungen ins Zentrum stellen: zum einen das mit externen Partnern zusammen durchgeführte multilinguale EU-Projekt „SprichWort“ und das Projekt „Usuelle Wortverbindungen“.
14. In *ellexiko* – einem Onlinewörterbuch zum Gegenwartsdeutschen – werden alle Artikel neu auf Basis des *ellexiko*-Korpus und gezielt für das elektronische Medium erarbeitet. Besonders interessant für den Kontext der Jahrestagung waren die Angabenbereiche der „Lexikalischen Mitspieler“ sowie der „Typischen Verwendungen“, die in ihrer Art singulär in *ellexiko* sind und anhand derer das Zusammenspiel von Korpusanalysen und lexikografischem Arbeiten exemplarisch demonstriert werden konnte.
15. Seit 2009 gibt es am IDS außerdem noch ein *ellexiko*-flankierendes Drittmittelprojekt mit dem Namen „Benutzeradaptive Zugänge und Vernetzungen in *ellexiko*“. Ein Schwerpunkt dieses Projektes ist die Benutzungsforschung in Bezug auf Online-Wörterbücher, da hier eine große Forschungslücke besteht. Neben der Benutzungsforschung geht es zum zweiten im Projekt um die Kodierung und Verwaltung sowie um Möglichkeiten der Darstellung von Vernetzungen im lexikografischen Kontext und zum Dritten um die automatische Gewinnung morphologischer Angaben für *ellexiko*.
16. Das letzte Projekt in der Reihe war das IDS-Projekt „OWID“, das „Online-Wortschatz-Informationssystem Deutsch“, in dem alle internetlexikografischen Arbeiten des IDS in einem Portal zusammengeführt werden. Momentan wird an einem kompletten Neuauftritt von OWID gearbeitet, den man sich auf der Projektmesse anschauen konnte und der noch in der ersten Hälfte des Jahres 2010 online freigeschaltet werden soll.

Aus der Sicht des unter (2) erwähnten Projekts bietet eine solche Projektmesse mehr Nutzen, als auf den ersten Blick sichtbar werden dürfte: Zahlreiche Kommentare und Diskussionen am Stand, die fast selbstverständlich zu erwarten waren, haben die Motivierung der Projektteilnehmer sehr positiv unterstützt und werden nützliche Kontaktaufnahmen mit Erfahrungsaustausch zur Folge haben. Zu den Diskutierenden zählten ja u. a. auch die anwesenden Experten der deutschsprachigen lexikografischen Branche. Daneben spielen folgende Positiva und Desiderate eine Rolle, die hier kurz aufgelistet werden sollen:

- a. Für einen Lexikografen, der sich mit neuen Informationstechnologien und deren möglichen Umsetzungen im Bereich der Lexikografie befassen muss, existiert in der theoretischen Literatur wenig, was die Fragen nach der neuen Datenvermittlung und vor allem der Datenstrukturierung im Hinblick auf die neue, durch die Informationsmedien veränderte soziale Handlung des Benutzers betrifft. Gemeint sind alle Benutzungssituationen, die in der theoretischen Literatur noch nicht analysiert worden sind, und die zum Beispiel bei der Erstellung eines Textes (so z. B. während Übersetzungs-

und Editionsarbeiten) sich auf eine Vielfalt von Internetquellen beziehen, assoziationsbedingte oder an mehrere Quellen zugleich gerichtete Anfragen mit eingeschlossen. Diese Projektmesse konnte dazu bereits zum Teil beitragen und neue Forschungen aufzeigen (vgl. oben unter 15).

- b. Unter den Tagungsbesuchern findet sich immer eine große Anzahl Linguisten, die in ihrem pädagogischen Alltag mit neuen, im Netz stehenden digitalen Nachschlagewerken arbeiten. Dabei sammeln sie wertvolle Erfahrungen, die sie jedoch bisher nicht publiziert haben oder nie publizieren werden. Ihrem Wissen steht ein zwar theoretisch begründetes, jedoch zum Teil auch weltfremdes Benutzer-Bild aus der metalexikografischen Literatur gegenüber, das man bei der Konzeptionsentwicklung eines lexikografischen Unternehmens vor dem inneren Auge konstruiert und das zur Beeinträchtigung des Machbaren führen kann. Eine Projektmesse bietet dabei die Gelegenheit, dass auch solche Kollegen zu Wort kommen, die ihre Erfahrung in der Wörterbuchbenutzung nicht publizieren, die aber den Projektleitern den authentischen

Bedarf unterschiedlicher Benutzerkreise sehr gut deutlich machen können. Nur durch solchen Austausch kann ein Nachschlagewerk oder eine Datenbank zu einem guten Werkzeug werden.

- c. Der Trend der heutigen Lexikografie, sprachliches Wissen mit dem Weltwissen im Rahmen der sich immer erweiternden Informationen darzustellen, wird fast alle beteiligten Kollegen herausfordern: Befindet sich nämlich ein Nachschlagewerk im Netz, wird es fast immer in Kombination mit anderen elektronischen Quellen konsultiert. Die Verlinkung der lexikalischen Datenbanken mit anderen Internet-Ressourcen bleibt darüber hinaus ein weiteres spannendes Thema.

Carolyn Müller-Spitzer ist wissenschaftliche Mitarbeiterin am Institut für Deutsche Sprache in Mannheim, Marie Vachková ist Leiterin der Lexikographischen Sektion am Institut für germanische Studien, Philosophische Fakultät der Karls-Universität in Prag.

Fotos: Annette Trabold.